TEST 3 **NAME**_____

Be clear. Be neat - no scratchouts accepted. Present your work in a folder. Make it easy to read and to follow.

1. Use the summary data (gas mileage per gallon) below to construct an One-Way ANOVA table and then answer the questions.

Truck	Average MPG	Variance of MPG	Sample size
Chevy	14.9	0.16	5
Dodge	14.4	0.07	5
Ford	14.5	0.20	5

(i) Does average gas mileage differ between the standard four-wheel drive pickup trucks made by Chevrolet, Dodge and Ford? Use $\alpha = .10$. State the hypotheses and provide a well-written conclusion.

(ii) Estimate the proportion of the overall variation in mpg that can be explained by the differing truck types?

(iii) Create a dataset that satisfies the table above,
then use MINITAB to verify your earlier calculations.
For example, to create the Chevy data:

```
MTB> random 5 c91
MTB> let k94=stdev(c91)
MTB> let k93=(.16)**.5
MTB> let c92=(k93/k94)*c91
MTB> let k95=mean(c92)
MTB> let c93=c92-k95+14.9
MTB> print c93
MTB> describe c93
MTB> copy c93 c1
MTB>erase c91-c93
MTB> erase k92-k95
```

(iv) Using the simulated data to provide appropriate (data and residual) plots to illustrate your conclusions and to check the validity of the one-way model assumptions.

(v) What sample sizes are needed for each truck type so that the power is at least .95 to detect a maximum difference in means of 0.50? Assume a commom standard deviation of 0.40.

2. A researcher conducted an experiment to compare the effects of three different insecticides on a variety of string beans. To obtain a sufficient amount of data, it was necessary to use four different plots of land. Since the plots had somewhat different soil fertility, drainage characteristics, and sheltering from winds, the researcher decided to conduct a randomized complete block design with the plots serving as blocks. Each plot was subdivided into three rows. A suitable distance was maintained between rows within a plot so that the insecticides could be confined to a particular row. Each row was planted with a 100 seeds and then maintained under the insecticide assigned to the row. The insecticides were randomly assigned to the rows within a plot so that each insecticide appeared in one row within all our plots. The response Y_{ij} of interest was the number of seedlings that emerged per row. The data and means are given in the table below.

		Plot			
Insecticide	1	2	3	4	Insecticide Mean
1	56	48	66	62	58
2	83	78	94	93	87
3	80	72	83	85	80
Plot Mean	73	66	81	80	75

(i) Fill in the blanks to complete the RCB model description:

$$Y_{ij} = \mu... + \_\_\_\_\_ \quad (1)$$

where $\mu...$ is a constant,

_____ are constants for the block (plot) effects,

_____ are constants for the treatment (insecticide) effects,

ϵ_{ij} are independent $N($ ____, $\sigma^2)$ random variables,

$i = 1, ..., 4$; $j = 1, ...,$ _____.

The responses Y_{ij} are independent and normally distributed

with mean $E\{Y_{ij}\} =$ _____

and variance $\sigma^2\{Y_{ij}\} =$ _____.

(ii) Provide point estimators for all nine parameters (unknown constants) in model 1.

(iii) Provide a plot of Number of Seedlings versus Insecticide Type (line-connected by plots). Discuss the model assumption of no interactions based on evidence from the plot.

(iv) Complete the ANOVA table.

Source of Variation	SS	df	MS	F^*	p-value
Blocks (plots)					
Treatments (insecticides)					
Error (unexplained)					
Total					

(v) Provide a residual plot that addresses the model assumption of error normality. Does the plot provide convincing evidence against this model assumptions? Discuss.

(vi) Provide two residual plots that address the model assumption of homoskedasticity. Do the plots provide convincing evidence against this model assumption? Discuss.

(vii) Perform a Tukey Test for Additivity (use α=.05).

(viii). Use $\alpha = .05$ and test for significant insecticide effects.

(ix) Estimate the proportion of the total response variability that is explained by the different insecticides used.

(x). Perform all 3 pairwise comparisons of the $\mu_{\cdot j}$ (insecticide means). To lessen the number of negative signs, let $D_1 = \mu_{\cdot 2} - \mu_{\cdot 1}$, $D_2 = \mu_{\cdot 3} - \mu_{\cdot 1}$, and $D_3 = \mu_{\cdot 2} - \mu_{\cdot 3}$. Use the Tukey procedure with a family confidence coefficient of .95. After constructing all intervals, note which means differ significantly.

(xi). Use $\alpha = .05$ and test for significant plot effects.

(xii) Estimate the proportion of the total response variability that is explained by the different plots used in the experiment.

(xiii) Explain why a randomized complete block design was appropriate for this experiment.